J. Soc. Cosmet. Chem., 44, 65–67 (January/February 1993)

Letters to the Editor

TO THE EDITOR:

The article entitled "Comparison of exaggerated and normal use techniques for assessing the mildness of personal cleansers" by B. H. Keswick *et al.* (*J. Soc. Cosmet. Chem.*, 43, 187–193, 1992) purports to demonstrate superior predictive ability and relevance of an arm wash protocol compared to a flex wash protocol. There are several pieces of data presented in the Keswick article that are contradicted by published data as well as being internally inconsistent. The lack of specific experimental details, such as the number of subjects, inclusion criteria, and description of grading scale(s) make comparison with the literature even more difficult. These deficiencies call into question the validity, not only of the conclusions, but of the actual data as well.

In a previously published study by Sharko *et al.* (1), two products with very similar formulations to products A and B described in the Keswick article were tested using an arm wash test. In the Sharko article, well-established instrumental assessment techniques were used in addition to the standard expert clinical evaluation such as that reported in the Keswick article. The product A analogue was significantly ($p < 0.05$) less irritating than the product B analogue and caused significantly less damage to the barrier as assessed by TEWL. There were directional differences ($0.05 < p < 0.1$) showing product B to cause a greater increase in blood flow when assessed by a Periflux Laser Doppler Velocimeter and greater redness when measured with a Minolta Chromameter. These results are self-supportive and consistent with our own flex wash tests and the flex wash reported in the Keswick article.

However, in the Keswick article, the authors claim that their arm wash test actually had results opposite to those reported in the two flex wash tests and the arm wash test mentioned above. We have conducted dozens of flex wash tests as well as less aggressive arm wash and face wash tests over many years, and have never observed such a cross-over.

The data from the in home-use test with the product pair reported in the Keswick article adds no value since there were no significant differences observed, except for a claimed significance for a very small difference in leg dryness. Even in this case, if any appropriate nonparametric analysis of the data were conducted, it is unlikely that the small difference reported on legs would be significant.

Since there is no question that the flex wash is the more aggressive procedure, it is quite surprising to observe that for products A, C, and D in the Keswick article, the "milder" arm wash protocol produced markedly greater erythema than the flex wash. This finding

is also not consistent with information already in the literature regarding clinical testing of surfactants (2).

Finally, the results of the appliance test reported in the Keswick article are very confusing and again inconsistent with the arm wash data presented in the same article. Although the product used was not identified, it may be inferred from the erythema score that the product is similar in irritation potential to product B. However, when using the towel, no irritation was observed—a lesser insult than that produced by the "milder" arm wash methodology. Keswick *et al.* do not attempt to explain these glaring inconsistencies. The description given of their consumer use model certainly cannot account for such inexplicable differences.

In summary, the results of the study of Keswick *et al.* do not support any known models of surfactant-induced irritation, and the proposed model for irritation produced in a flex wash is simplistic and inaccurate. Not only are the data internally inconsistent, but they are contradicted by published information. These major discrepancies must be addressed.

REFERENCES

(1) P. T. Sharko *et al.*, Arm wash with instrumental evaluation—A sensitive technique for differentiating the irritation potential of personal washing products. *J. Dermatol. Clin. Eval. Soc.*, **2**, 19–27 (1991).
(2) P. J. Frosch, "Irritancy of Soaps and Detergent Bars, in *Principles of Cosmetics for the Dermatologist*, P. Frost and S. N. Horwitz, Eds. (C.V. Mosby Co., St. Louis, 1982), pp. 5–12.

Richard I. Murahata
Unilever Research U.S.

Gary L. Grove
KGL, Inc.

TO THE EDITOR:

This responds to comments made regarding our article, "Comparison of exaggerated and normal use techniques for assessing the mildness of personal cleansers," which recently appeared in your journal. Specifically, this addresses criticisms that the article provided insufficient experimental detail and that the results presented are inconsistent.

As we noted in our article, the forearm and flex wash methods examined were based on methods previously published in this journal (1,2), and the procedural details requested are available in the cited works. Neither the reviewers nor the editor deemed it necessary to repeat these details in our article. For those who may not have read the original articles, the forearm and flex wash studies were graded on seven-point and four-point scales, respectively. The forearm wash studies used 30–40 subjects for each product comparison, the flex wash implement study used 13 subjects for each treatment comparison, and the flex wash studies used 40 subjects for each product comparison. These sample sizes met or exceeded the sample sizes specified in the original papers. The home-use studies were scored using the same grading scales as the forearm wash studies, with between 80 and 90 subjects randomly assigned to use each test product.

The point made by our article is that exaggerated wash procedures differ in their ability to predict skin effects induced by personal cleansing products under conditions of actual consumer use. Specifically, our results indicate that the forearm wash method we used yields results that are more consistent with the skin effects experienced by consumers under home-use conditions than does the flex wash method. The correspondents selectively criticize these findings, not by presenting data that show the flex wash method to be predictive of skin effects observed under conditions of consumer use, but on the basis of results generated by a different forearm wash method and traditional theories of surfactant irritation. Neither of these arguments establishes a correlation between the product mildness predicted by the flex wash method and that observed when the product is used by consumers. In fact, the forearm wash results cited by the correspondents failed to differentiate the products in terms of their ability to induce dryness, a product attribute that is very noticeable to consumers. Further, the traditional arguments presented by the correspondents do not hold for systems employing new technologies, specifically personal cleansing bars incorporating cationic polymers. The skin protectant properties of these materials are well-documented in the literature, and we have performed numerous studies that have demonstrated the clinical mildness benefits of this technology over simple, surfactant-based systems.

The flex wash method differs from the arm wash method we used in that it involves a greater amount of mechanical manipulation (of the skin) during the washing process. The correspondents acknowledge that the flex wash method is an aggressive procedure, and the results from our implement experiment show that applying product with a sponge causes greater disruption of the stratum corneum barrier than does applying product with a towel. Overly aggressive wash methods may fail to demonstrate positive skin effects introduced by new technologies and, thus, fail to predict skin effects observed by consumers, because of the greater amount of barrier damage they induce. We believe that the disagreement noted by the correspondents casts doubt, not on the validity of our data, but on the ability of overly aggressive wash methods and traditional surfactant-irritation theories to account for skin mildness benefits afforded by new technologies.

A second criticism offered by the correspondents is a lack of internal consistency of the data presented for the various products. This criticism is based on erroneous comparisons of absolute erythema scores generated by different wash protocols and scored on different grading scales and, hence, is unfounded.

REFERENCES

(1) M. F. Lukacovic, F. E. Dunlap, S. E. Michaels, M. O. Visscher, and D. D. Watson, Forearm wash test to evaluate the mildness of cleansing products, *J. Soc. Cosmet. Chem.*, **39**, 355–366 (1988).

(2) D. D. Strube, S. W. Koontz, R. I. Murahata, and R. F. Theiler, The flex wash test: A method for evaluating the mildness of personal washing products, *J. Soc. Cosmet. Chem.*, **40**, 297–306 (1989).

B. H. Keswick
K. D. Ertel
M. O. Visscher